

SEQUENCE LISTING

<110> ROSENBERG, Robert D
 SHWORAK, Nicholas W
 LIU, Jian
 FRITZE, Linda M. S.
 SCHWARTZ, John J
 ZHANG, Lijuan

<120> HEPARAN SULFATE D-GLUCOSAMINYL 3-O-SULFOTRANSFERASES,
 AND USES THEREFOR

<130> MIT-087

<140>

<141>

<150> WO PCT/US98/22597

<151> 1998-10-23

<150> USSN 60/065,437

<151> 1997-10-31

<150> USSN 60/062,762

<151> 1997-10-24

<160> 16

<170> PatentIn Ver. 2.0

<210> 1

<211> 1685

<212> DNA

<213> Mus musculus

<220>

<221> CDS

<222> (323)..(1255)

<223> mouse 3-OST-1

<400> 1

tgcattgcaa tgtgaagtgt tctgaataa acctgcttga agaaggacaa cgtggtgttg 60

cgtctttcct gctggtcggg gtggaataga cacctcccct ttttaacttg ggtgacctca 120

tgaacataaa agaacttaaa ggtagcaagc catggactta aagtaggctg accttgaact 180

cagagatcct cttggcaatg tctctggaga ttaaagtaat tggcaactgg agatactcat 240

gttccagtaa tcaagaggga gccttgctgc tacttcatga tccaggcgcg tgtggcccag 300

tgaagtcct gagctgtaca gc atg acc ttg ctg ctc ctg ggt gcg gtg ctg 352

Met Thr Leu Leu Leu Leu Gly Ala Val Leu

1

5

10

ctg gtg gcc cag ccc cag ctt gtg cat tcc cac ccg gct gct cct ggc 400

Leu Val Ala Gln Pro Gln Leu Val His Ser His Pro Ala Ala Pro Gly

15															20					25					
ccg	ggg	ctc	aaa	cag	cag	gag	ctt	ctg	agg	aag	gtg	att	att	ctc	cca	448									
Pro	Gly	Leu	Lys	Gln	Gln	Glu	Leu	Leu	Arg	Lys	Val	Ile	Ile	Leu	Pro										
			30				35				40														
gag	gac	acc	gga	gaa	ggc	aca	gca	tcc	aat	ggt	tcc	aca	cag	cag	ctg	496									
Glu	Asp	Thr	Gly	Glu	Gly	Thr	Ala	Ser	Asn	Gly	Ser	Thr	Gln	Gln	Leu										
			45				50				55														
cca	cag	acc	atc	atc	att	ggg	gtg	cgc	aag	ggt	ggt	acc	cga	gcc	ctg	544									
Pro	Gln	Thr	Ile	Ile	Ile	Gly	Val	Arg	Lys	Gly	Gly	Thr	Arg	Ala	Leu										
			60				65				70														
cta	gag	atg	ctc	agc	ctg	cat	cct	gat	gtt	gct	gca	gct	gaa	aac	gag	592									
Leu	Glu	Met	Leu	Ser	Leu	His	Pro	Asp	Val	Ala	Ala	Ala	Glu	Asn	Glu										
75				80				85				90													
gtc	cat	ttc	ttt	gac	tgg	gag	gag	cat	tac	agc	caa	ggc	ctg	ggc	tgg	640									
Val	His	Phe	Phe	Asp	Trp	Glu	Glu	His	Tyr	Ser	Gln	Gly	Leu	Gly	Trp										
			95				100				105														
tac	ctc	acc	cag	atg	ccc	ttc	tcc	tcc	cct	cac	cag	ctc	acc	gtg	gag	688									
Tyr	Leu	Thr	Gln	Met	Pro	Phe	Ser	Ser	Pro	His	Gln	Leu	Thr	Val	Glu										
			110				115				120														
aag	aca	ccc	gcc	tat	ttc	act	tgc	ccc	aaa	gtg	cct	gag	aga	atc	cac	736									
Lys	Thr	Pro	Ala	Tyr	Phe	Thr	Ser	Pro	Lys	Val	Pro	Glu	Arg	Ile	His										
			125				130				135														
agc	atg	aac	ccc	acc	atc	cgc	ctg	ctg	ctt	atc	ctg	agg	gac	cca	tca	784									
Ser	Met	Asn	Pro	Thr	Ile	Arg	Leu	Leu	Leu	Ile	Leu	Arg	Asp	Pro	Ser										
			140				145				150														
gag	cgc	gtg	ctg	tcc	gac	tac	acc	cag	gtg	ttg	tac	aac	cac	ctt	cag	832									
Glu	Arg	Val	Leu	Ser	Asp	Tyr	Thr	Gln	Val	Leu	Tyr	Asn	His	Leu	Gln										
155				160				165				170													
aag	cac	aag	ccc	tat	cca	ccc	att	gag	gac	ctc	cta	atg	cgg	gac	ggt	880									
Lys	His	Lys	Pro	Tyr	Pro	Pro	Ile	Glu	Asp	Leu	Leu	Met	Arg	Asp	Gly										
			175				180				185														
cgg	ctg	aac	ctg	gac	tac	aag	gct	ctc	aac	cgc	agc	ctg	tac	cat	gca	928									
Arg	Leu	Asn	Leu	Asp	Tyr	Lys	Ala	Leu	Asn	Arg	Ser	Leu	Tyr	His	Ala										
			190				195				200														
cac	atg	ctg	aac	tgg	ctg	cgt	ttt	ttc	ccg	ttg	ggc	cac	atc	cac	att	976									
His	Met	Leu	Asn	Trp	Leu	Arg	Phe	Phe	Pro	Leu	Gly	His	Ile	His	Ile										
			205				210				215														
gtg	gat	ggc	gac	cgc	ctc	atc	aga	gac	cct	ttc	cct	gag	atc	cag	aag	1024									
Val	Asp	Gly	Asp	Arg	Leu	Ile	Arg	Asp	Pro	Phe	Pro	Glu	Ile	Gln	Lys										
			220				225				230														
gtc	gaa	aga	ttc	ctg	aag	ctt	tct	cca	cag	atc	aac	gcc	tgc	aac	ttc	1072									
Val	Glu	Arg	Phe	Leu	Lys	Leu	Ser	Pro	Gln	Ile	Asn	Ala	Ser	Asn	Phe										
235				240				245				250													

tac ttt aac aaa acc aag ggc ttc tac tgc ctg cgg gac agt ggc aag 1120
 Tyr Phe Asn Lys Thr Lys Gly Phe Tyr Cys Leu Arg Asp Ser Gly Lys
 255 260 265

gac cgc tgc tta cac gag tcc aaa ggc cgg gcg cac ccc cag gtg gat 1168
 Asp Arg Cys Leu His Glu Ser Lys Gly Arg Ala His Pro Gln Val Asp
 270 275 280

ccc aaa cta ctt gat aaa ctg cac gaa tac ttt cat gag cca aat aag 1216
 Pro Lys Leu Leu Asp Lys Leu His Glu Tyr Phe His Glu Pro Asn Lys
 285 290 295

aaa ttt ttc aag ctc gtg ggc aga aca ttc gac tgg cac tgatttgccg 1265
 Lys Phe Phe Lys Leu Val Gly Arg Thr Phe Asp Trp His
 300 305 310

tctcctaggc tcgggacttt tctgttggt aacttctggt gtacatctga aggggggagg 1325

aaaataattt taaaaggca tttaagctat aatttatttg taaaaccac aaatgacttc 1385

tgtacagtat tagattcaca gttgccatat atagtagtta tatttttcta cttgttaa 1445

ggagggcggt ttgtattggt tttcatgggt gttaacattg tgtatatgtc tctataatat 1505

gaaggaactt aactattgca ctgaaaaaat aagagatttt ttttttctgg agacctcttt 1565

ttttgttggt gttgttttaa atataattaa cctgcctcca atccaaaata gctctttggt 1625

ttcacctcct tgtcaaactc ataacttttt tctgcttaaa aaatttattg gtattatgga 1685

<210> 2
 <211> 311
 <212> PRT
 <213> Mus musculus

<400> 2
 Met Thr Leu Leu Leu Leu Gly Ala Val Leu Leu Val Ala Gln Pro Gln
 1 5 10 15

Leu Val His Ser His Pro Ala Ala Pro Gly Pro Gly Leu Lys Gln Gln
 20 25 30

Glu Leu Leu Arg Lys Val Ile Ile Leu Pro Glu Asp Thr Gly Glu Gly
 35 40 45

Thr Ala Ser Asn Gly Ser Thr Gln Gln Leu Pro Gln Thr Ile Ile Ile
 50 55 60

Gly Val Arg Lys Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Ser Leu
 65 70 75 80

His Pro Asp Val Ala Ala Ala Glu Asn Glu Val His Phe Phe Asp Trp
 85 90 95

Glu Glu His Tyr Ser Gln Gly Leu Gly Trp Tyr Leu Thr Gln Met Pro

100	105	110
Phe Ser Ser Pro His Gln Leu Thr Val Glu Lys Thr Pro Ala Tyr Phe 115 120 125		
Thr Ser Pro Lys Val Pro Glu Arg Ile His Ser Met Asn Pro Thr Ile 130 135 140		
Arg Leu Leu Leu Ile Leu Arg Asp Pro Ser Glu Arg Val Leu Ser Asp 145 150 155 160		
Tyr Thr Gln Val Leu Tyr Asn His Leu Gln Lys His Lys Pro Tyr Pro 165 170 175		
Pro Ile Glu Asp Leu Leu Met Arg Asp Gly Arg Leu Asn Leu Asp Tyr 180 185 190		
Lys Ala Leu Asn Arg Ser Leu Tyr His Ala His Met Leu Asn Trp Leu 195 200 205		
Arg Phe Phe Pro Leu Gly His Ile His Ile Val Asp Gly Asp Arg Leu 210 215 220		
Ile Arg Asp Pro Phe Pro Glu Ile Gln Lys Val Glu Arg Phe Leu Lys 225 230 235 240		
Leu Ser Pro Gln Ile Asn Ala Ser Asn Phe Tyr Phe Asn Lys Thr Lys 245 250 255		
Gly Phe Tyr Cys Leu Arg Asp Ser Gly Lys Asp Arg Cys Leu His Glu 260 265 270		
Ser Lys Gly Arg Ala His Pro Gln Val Asp Pro Lys Leu Leu Asp Lys 275 280 285		
Leu His Glu Tyr Phe His Glu Pro Asn Lys Lys Phe Phe Lys Leu Val 290 295 300		
Gly Arg Thr Phe Asp Trp His 305 310		

<210> 3

<211> 1305

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (119)..(1039)

<223> human 3-OST-1

<400> 3

cgcggtctcag taattgaagg cctgaaacgc ccatgtgcc a ctgactagga ggcttcctg 60

ctgcggcact tcatgaccca ggggcgcgcg gccacgtgaa gccaccgtgg tgtccagc 118

atg gcc gcg ctg ctc ctg ggc gcg gtg ctg ctg gtg gcc cag ccc cag	166
Met Ala Ala Leu Leu Leu Gly Ala Val Leu Leu Val Ala Gln Pro Gln	
1 5 10 15	
cta gtg cct tcc cgc ccc gcc gag cta ggc cag cag gag ctt ctg cgg	214
Leu Val Pro Ser Arg Pro Ala Glu Leu Gly Gln Gln Glu Leu Leu Arg	
20 25 30	
aaa gcg ggg acc ctc cag gat gac gtc cgc gat ggc gtg gcc cca aac	262
Lys Ala Gly Thr Leu Gln Asp Asp Val Arg Asp Gly Val Ala Pro Asn	
35 40 45	
ggc tct gcc cag cag ttg ccg cag acc atc atc atc ggc gtg cgc aag	310
Gly Ser Ala Gln Gln Leu Pro Gln Thr Ile Ile Ile Gly Val Arg Lys	
50 55 60	
ggc gcc acg cgc gca ctg ctg gag atg ctc agc ctg cac ccc gac gtg	358
Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Ser Leu His Pro Asp Val	
65 70 75 80	
gcg gcc gcg gag aac gag gtc cac ttc ttc gac tgg gag gag cat tac	406
Ala Ala Ala Glu Asn Glu Val His Phe Phe Asp Trp Glu Glu His Tyr	
85 90 95	
agc cac gcc ttg ggc tgg tac ctc agc cag atg ccc ttc tcc tgg cca	454
Ser His Gly Leu Gly Trp Tyr Leu Ser Gln Met Pro Phe Ser Trp Pro	
100 105 110	
cac cag ctc aca gtg gag aag acc ccc gcg tat ttc acg tcg ccc aaa	502
His Gln Leu Thr Val Glu Lys Thr Pro Ala Tyr Phe Thr Ser Pro Lys	
115 120 125	
gtg cct gag cga gtc tac agc atg aac ccg tcc atc cgg ctg ctg ctc	550
Val Pro Glu Arg Val Tyr Ser Met Asn Pro Ser Ile Arg Leu Leu Leu	
130 135 140	
atc ctg cga gac ccg tcg gag cgc gtg cta tct gac tac acc caa gtg	598
Ile Leu Arg Asp Pro Ser Glu Arg Val Leu Ser Asp Tyr Thr Gln Val	
145 150 155 160	
ttc tac aac cac atg cag aag cac aag ccc tac ccg tcc atc gag gag	646
Phe Tyr Asn His Met Gln Lys His Lys Pro Tyr Pro Ser Ile Glu Glu	
165 170 175	
ttc ctg gtg cgc gat ggc agg ctc aat gtg gac tac aag gcc ctc aac	694
Phe Leu Val Arg Asp Gly Arg Leu Asn Val Asp Tyr Lys Ala Leu Asn	
180 185 190	
cgc agc ctc tac cac gtg cac atg cag aac tgg ctg cgc ttt ttc ccg	742
Arg Ser Leu Tyr His Val His Met Gln Asn Trp Leu Arg Phe Phe Pro	
195 200 205	
ctg cgc cac atc cac att gtg gac ggc gac cgc ctc atc agg gac ccc	790
Leu Arg His Ile His Ile Val Asp Gly Asp Arg Leu Ile Arg Asp Pro	
210 215 220	
ttc cct gag atc caa aag gtc gag agg ttc cta aag ctg tcg ccg cag	838

Phe Pro Glu Ile Gln Lys Val Glu Arg Phe Leu Lys Leu Ser Pro Gln
 225 230 235 240
 atc aat gct tcg aac ttc tac ttt aac aaa acc aag ggc ttt tac tgc 886
 Ile Asn Ala Ser Asn Phe Tyr Phe Asn Lys Thr Lys Gly Phe Tyr Cys
 245 250 255
 ctg cgg gac agc ggc cgg gac cgc tgc tta cat gag tcc aaa ggc cgg 934
 Leu Arg Asp Ser Gly Arg Asp Arg Cys Leu His Glu Ser Lys Gly Arg
 260 265 270
 gcg cac ccc caa gtc gat ccc aaa cta ctc aat aaa ctg cac gaa tat 982
 Ala His Pro Gln Val Asp Pro Lys Leu Leu Asn Lys Leu His Glu Tyr
 275 280 285
 ttt cat gag cca aat aag aag ttc ttc gag ctt gtt ggc aga aca ttt 1030
 Phe His Glu Pro Asn Lys Lys Phe Phe Glu Leu Val Gly Arg Thr Phe
 290 295 300
 gac tgg cac tgatttgcaa taagctaagc tcagaaactt tcctactgta 1079
 Asp Trp His
 305
 agttctggtg tacatctgag gggaaaaaga attttaaaaa agcatttaag gtataattta 1139
 tttgtaaaat ccataaagta cttctgtaca gtattagatt cacaattgcc atatatacta 1199
 gttatatttt tctacttggt aaatggaggg cattttgtat tgtttttcat ggttggttaac 1259
 attgtgtaat atgtctctat atgaaggaac taaactattt cactga 1305

<210> 4
 <211> 307
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Ala Ala Leu Leu Leu Gly Ala Val Leu Leu Val Ala Gln Pro Gln
 1 5 10 15
 Leu Val Pro Ser Arg Pro Ala Glu Leu Gly Gln Gln Glu Leu Leu Arg
 20 25 30
 Lys Ala Gly Thr Leu Gln Asp Asp Val Arg Asp Gly Val Ala Pro Asn
 35 40 45
 Gly Ser Ala Gln Gln Leu Pro Gln Thr Ile Ile Ile Gly Val Arg Lys
 50 55 60
 Gly Gly Thr Arg Ala Leu Leu Glu Met Leu Ser Leu His Pro Asp Val
 65 70 75 80
 Ala Ala Ala Glu Asn Glu Val His Phe Phe Asp Trp Glu Glu His Tyr
 85 90 95
 Ser His Gly Leu Gly Trp Tyr Leu Ser Gln Met Pro Phe Ser Trp Pro

100	105	110
His Gln Leu Thr Val Glu Lys Thr Pro Ala Tyr Phe Thr Ser Pro Lys 115 120 125		
Val Pro Glu Arg Val Tyr Ser Met Asn Pro Ser Ile Arg Leu Leu Leu 130 135 140		
Ile Leu Arg Asp Pro Ser Glu Arg Val Leu Ser Asp Tyr Thr Gln Val 145 150 155 160		
Phe Tyr Asn His Met Gln Lys His Lys Pro Tyr Pro Ser Ile Glu Glu 165 170 175		
Phe Leu Val Arg Asp Gly Arg Leu Asn Val Asp Tyr Lys Ala Leu Asn 180 185 190		
Arg Ser Leu Tyr His Val His Met Gln Asn Trp Leu Arg Phe Phe Pro 195 200 205		
Leu Arg His Ile His Ile Val Asp Gly Asp Arg Leu Ile Arg Asp Pro 210 215 220		
Phe Pro Glu Ile Gln Lys Val Glu Arg Phe Leu Lys Leu Ser Pro Gln 225 230 235 240		
Ile Asn Ala Ser Asn Phe Tyr Phe Asn Lys Thr Lys Gly Phe Tyr Cys 245 250 255		
Leu Arg Asp Ser Gly Arg Asp Arg Cys Leu His Glu Ser Lys Gly Arg 260 265 270		
Ala His Pro Gln Val Asp Pro Lys Leu Leu Asn Lys Leu His Glu Tyr 275 280 285		
Phe His Glu Pro Asn Lys Lys Phe Phe Glu Leu Val Gly Arg Thr Phe 290 295 300		
Asp Trp His 305		

<210> 5
 <211> 1951
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (73)..(1173)
 <223> human 3-OST-2

<400> 5
 cgcaggggcca cagcagctca gccgcgggtg cccctcgga aaccatgacc cccggcgcg 60

gcccatggag cc atg gcc tat agg gtc ctg ggc cgc gcg ggg cca cct cag 111
 Met Ala Tyr Arg Val Leu Gly Arg Ala Gly Pro Pro Gln

1					5					10							
ccg	cgg	agg	gcg	cgc	agg	ctg	ctc	ttc	gcc	ttc	acg	ctc	tcg	ctc	tcc	159	
Pro	Arg	Arg	Ala	Arg	Arg	Leu	Leu	Phe	Ala	Phe	Thr	Leu	Ser	Leu	Ser		
15					20					25							
tgc	act	tac	ctg	tgt	tac	agc	ttc	ctg	tgc	tgc	tgc	gac	gac	ctg	ggg	207	
Cys	Thr	Tyr	Leu	Cys	Tyr	Ser	Phe	Leu	Cys	Cys	Cys	Asp	Asp	Leu	Gly		
30					35					40					45		
cgg	agc	cgc	ctc	ctc	ggc	gcg	cct	cgc	tgc	ctc	cgc	ggc	ccc	agc	gcg	255	
Arg	Ser	Arg	Leu	Leu	Gly	Ala	Pro	Arg	Cys	Leu	Arg	Gly	Pro	Ser	Ala		
50					55					60							
ggc	ggc	cag	aaa	ctt	ctc	cag	aag	tcc	cgc	ccc	tgt	gat	ccc	tcc	ggg	303	
Gly	Gly	Gln	Lys	Leu	Leu	Gln	Lys	Ser	Arg	Pro	Cys	Asp	Pro	Ser	Gly		
65					70					75							
ccg	acg	ccc	agc	gag	ccc	agc	gct	ccc	agc	gcg	ccc	gcc	gcc	gcc	gtg	351	
Pro	Thr	Pro	Ser	Glu	Pro	Ser	Ala	Pro	Ser	Ala	Pro	Ala	Ala	Ala	Val		
80					85					90							
ccc	gcc	cct	cgc	ctc	tcc	ggg	tcc	aac	cac	tcc	ggc	tca	ccc	aag	ctg	399	
Pro	Ala	Pro	Arg	Leu	Ser	Gly	Ser	Asn	His	Ser	Gly	Ser	Pro	Lys	Leu		
95					100					105							
ggg	acc	aag	cgg	ttg	ccc	caa	gcc	ctc	att	gtg	ggc	gtg	aag	aag	ggg	447	
Gly	Thr	Lys	Arg	Leu	Pro	Gln	Ala	Leu	Ile	Val	Gly	Val	Lys	Lys	Gly		
110					115					120					125		
ggc	acc	cgg	gcc	gtg	ctg	gag	ttt	atc	cga	gta	cac	ccg	gac	gtg	cgg	495	
Gly	Thr	Arg	Ala	Val	Leu	Glu	Phe	Ile	Arg	Val	His	Pro	Asp	Val	Arg		
130					135					140							
gcc	ttg	ggc	acg	gaa	ccc	cac	ttc	ttt	gac	agg	aac	tac	ggc	cgc	ggg	543	
Ala	Leu	Gly	Thr	Glu	Pro	His	Phe	Phe	Asp	Arg	Asn	Tyr	Gly	Arg	Gly		
145					150					155							
ctg	gat	tgg	tac	agg	agc	ctg	atg	ccc	agg	acc	ctc	gag	agc	cag	atc	591	
Leu	Asp	Trp	Tyr	Arg	Ser	Leu	Met	Pro	Arg	Thr	Leu	Glu	Ser	Gln	Ile		
160					165					170							
acg	ctg	gag	aag	acg	ccc	agc	tac	ttt	gtc	act	caa	gag	gct	cct	cga	639	
Thr	Leu	Glu	Lys	Thr	Pro	Ser	Tyr	Phe	Val	Thr	Gln	Glu	Ala	Pro	Arg		
175					180					185							
cgc	atc	ttc	aac	atg	tcc	cga	gac	acc	aag	ctg	atc	gtg	gtt	gtg	cgg	687	
Arg	Ile	Phe	Asn	Met	Ser	Arg	Asp	Thr	Lys	Leu	Ile	Val	Val	Val	Arg		
190					195					200					205		
aac	cct	gtg	acc	cgt	gcc	atc	tct	gat	tac	acg	cag	aca	ctc	tcc	aag	735	
Asn	Pro	Val	Thr	Arg	Ala	Ile	Ser	Asp	Tyr	Thr	Gln	Thr	Leu	Ser	Lys		
210					215					220							
aag	ccc	gac	atc	ccg	acc	ttt	gag	ggc	ctc	tcc	ttc	cgc	aac	cgc	acc	783	
Lys	Pro	Asp	Ile	Pro	Thr	Phe	Glu	Gly	Leu	Ser	Phe	Arg	Asn	Arg	Thr		
225					230					235							

ctg ggc ctg gtg gac gtg tgc tgg aac gcc atc cgc atc ggc atg tac	831
Leu Gly Leu Val Asp Val Ser Trp Asn Ala Ile Arg Ile Gly Met Tyr	
240 245 250	
gtg ctg cac ctg gag agc tgg ctg cag tac ttc ccg cta gct cag att	879
Val Leu His Leu Glu Ser Trp Leu Gln Tyr Phe Pro Leu Ala Gln Ile	
255 260 265	
cac ttc gtc agt ggc gag cga ctc atc act gac ccg gcc ggc gag atg	927
His Phe Val Ser Gly Glu Arg Leu Ile Thr Asp Pro Ala Gly Glu Met	
270 275 280 285	
ggg cga gtc cag gac ttc ctg ggc att aag aga ttc atc acg gac aag	975
Gly Arg Val Gln Asp Phe Leu Gly Ile Lys Arg Phe Ile Thr Asp Lys	
290 295 300	
cac ttc tat ttc aac aag acc aaa gga ttc cct tgc ttg aaa aaa aca	1023
His Phe Tyr Phe Asn Lys Thr Lys Gly Phe Pro Cys Leu Lys Lys Thr	
305 310 315	
gaa tgc agc ctc ctg cct cga tgc ttg ggc aaa tca aaa ggc aga act	1071
Glu Ser Ser Leu Leu Pro Arg Cys Leu Gly Lys Ser Lys Gly Arg Thr	
320 325 330	
cat gta cag att gat cct gaa gtg ata gac cag ctc cga gaa ttt tat	1119
His Val Gln Ile Asp Pro Glu Val Ile Asp Gln Leu Arg Glu Phe Tyr	
335 340 345	
aga ccg tat aat atc aaa ttt tat gaa acc gtt ggc cag gac ttc agg	1167
Arg Pro Tyr Asn Ile Lys Phe Tyr Glu Thr Val Gly Gln Asp Phe Arg	
350 355 360 365	
tgg gaa taagcccacg aaaggaaagg gctctcaagg gctcttctgc tcatctcttc	1223
Trp Glu	
cgtagagattt gctcccagac cctcttatct cctccaaca aacctgggt ccagccccct	1283
ttcccaactt gagttgcac atcttggaac caggaagccc agctaaagcc aagagaccag	1343
agagtctctg ccactagttt tcatcagtct gttcaagcaa agttgatctg ctctggcac	1403
gtccagtaaa ttccagaatc attctccttt ctgccataa agggccttgg agaattgctt	1463
taagaagagt gaatgtcca atgatgatag atattataag cgacgatggg tctgttgcta	1523
tgaacacagc agtcgggtccc tgctattgtc caccaggag tggccttggt aattccaagt	1583
ggcatgtatc ttccctctga gcttcatttc ttcaagatgc tctgggtggg gggatgggag	1643
accatcctca gccctcctca gaccttatca attcattgag agattgcaaa gctgaaagca	1703
cctccggcca ctctggggag acagaccctt tgggtgatgaa ataaaccagt gacttcagag	1763
cctatgggtct caactgtgct tgaaaaacac tgtctctgaa aacaactttg tgattctccc	1823
tgtccctgt ggacaaaagc acataattct gctgttacgg gtactttgct catacgagct	1883

ttcatgttca gcatgcaatg gaatcatgct tgtccatgtg aaataaatat ggctctctcg 1943

tgtcctta

1951

<210> 6

<211> 367

<212> PRT

<213> Homo sapiens

<400> 6

Met Ala Tyr Arg Val Leu Gly Arg Ala Gly Pro Pro Gln Pro Arg Arg
1 5 10 15

Ala Arg Arg Leu Leu Phe Ala Phe Thr Leu Ser Leu Ser Cys Thr Tyr
20 25 30

Leu Cys Tyr Ser Phe Leu Cys Cys Cys Asp Asp Leu Gly Arg Ser Arg
35 40 45

Leu Leu Gly Ala Pro Arg Cys Leu Arg Gly Pro Ser Ala Gly Gly Gln
50 55 60

Lys Leu Leu Gln Lys Ser Arg Pro Cys Asp Pro Ser Gly Pro Thr Pro
65 70 75 80

Ser Glu Pro Ser Ala Pro Ser Ala Pro Ala Ala Val Pro Ala Pro
85 90 95

Arg Leu Ser Gly Ser Asn His Ser Gly Ser Pro Lys Leu Gly Thr Lys
100 105 110

Arg Leu Pro Gln Ala Leu Ile Val Gly Val Lys Lys Gly Gly Thr Arg
115 120 125

Ala Val Leu Glu Phe Ile Arg Val His Pro Asp Val Arg Ala Leu Gly
130 135 140

Thr Glu Pro His Phe Phe Asp Arg Asn Tyr Gly Arg Gly Leu Asp Trp
145 150 155 160

Tyr Arg Ser Leu Met Pro Arg Thr Leu Glu Ser Gln Ile Thr Leu Glu
165 170 175

Lys Thr Pro Ser Tyr Phe Val Thr Gln Glu Ala Pro Arg Arg Ile Phe
180 185 190

Asn Met Ser Arg Asp Thr Lys Leu Ile Val Val Val Arg Asn Pro Val
195 200 205

Thr Arg Ala Ile Ser Asp Tyr Thr Gln Thr Leu Ser Lys Lys Pro Asp
210 215 220

Ile Pro Thr Phe Glu Gly Leu Ser Phe Arg Asn Arg Thr Leu Gly Leu
225 230 235 240

Val	Asp	Val	Ser	Trp	Asn	Ala	Ile	Arg	Ile	Gly	Met	Tyr	Val	Leu	His
				245					250					255	
Leu	Glu	Ser	Trp	Leu	Gln	Tyr	Phe	Pro	Leu	Ala	Gln	Ile	His	Phe	Val
				260				265					270		
Ser	Gly	Glu	Arg	Leu	Ile	Thr	Asp	Pro	Ala	Gly	Glu	Met	Gly	Arg	Val
		275					280					285			
Gln	Asp	Phe	Leu	Gly	Ile	Lys	Arg	Phe	Ile	Thr	Asp	Lys	His	Phe	Tyr
	290					295					300				
Phe	Asn	Lys	Thr	Lys	Gly	Phe	Pro	Cys	Leu	Lys	Lys	Thr	Glu	Ser	Ser
305					310					315					320
Leu	Leu	Pro	Arg	Cys	Leu	Gly	Lys	Ser	Lys	Gly	Arg	Thr	His	Val	Gln
				325					330					335	
Ile	Asp	Pro	Glu	Val	Ile	Asp	Gln	Leu	Arg	Glu	Phe	Tyr	Arg	Pro	Tyr
		340						345					350		
Asn	Ile	Lys	Phe	Tyr	Glu	Thr	Val	Gly	Gln	Asp	Phe	Arg	Trp	Glu	
		355					360					365			

<210> 7

<211> 2314

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (799)..(2016)

<223> human 3-OST-3A

<400> 7

```

cagcggcgcc ccaggaggca gccggtgagc gcctgcgagc agagtggcgg gggccgctga 60
cagggtcccg gcagcccagc ccagcccagc cagcgggctc acagggtggg tccaagagca 120
gtttggagca acccggcgct acggagaggg gtggacggct ctgcacgggc ctctgtctc 180
ccgctcgggc agagggactc ggggggacct cgctccttgg ccgagagaac ctgaactcgg 240
gcggagagaa cgcgcccagg cgggcaaggg gaccagagaa agccggggct ggaagtcact 300
gtcgctcgcc actgtctgga gcgcacggag cgcagaggcc cggcagccgc gcgtgcctc 360
ccggggaccg agccagtgat gcaggatcgc tgagcggaga tccgcgccga gaagtctctc 420
ggggccgggg ctgagacgca cgccttcgac accgctgcc aagccccgat tccggcgact 480
cttgccggga accgaggggc caaggctgcc ccaagctcag gacttgggcg agtctaagac 540
gatggtttct taagcacgga cccgcgttcc ccttcccgcc cctcgcactg gaggcagggg 600
tctgcgcgg ggcccccggg attccgtttc cccgcggagc cccggccgct gcctcccggg 660

```

acagttcgcacggccacaggggcgcacgggcgatgtggcctccgtccagcgcgctggccccg																720
ccgggggggatgctctggcacctgtcgggggtccaggccctagcatggccgggcgcgttgccccg																780
acgtcgcctccggctaggatg gcc cct ccg ggc ccg gcc agt gcc ctc tcc																831
Met Ala Pro Pro Gly Pro Ala Ser Ala Leu Ser																
1510																
acc tcg gcc gag ccg ctg tcc cgc agc atc ttc cgg aag ttc ttg ctg																879
Thr Ser Ala Glu Pro Leu Ser Arg Ser Ile Phe Arg Lys Phe Leu Leu																
152025																
atg ctc tgc tcc ctg ctc acg tcc ctt tac gtc ttc tac tgc ctg gcc																927
Met Leu Cys Ser Leu Leu Thr Ser Leu Tyr Val Phe Tyr Cys Leu Ala																
303540																
gag cgc tgc cag acc ctg tcc ggc ccc gtc gtg ggg ctg tcc ggc ggc																975
Glu Arg Cys Gln Thr Leu Ser Gly Pro Val Val Gly Leu Ser Gly Gly																
455055																
ggc gag gag gcg ggg gcc cct ggt ggc ggc gtc ctg gcc gga ggc ccg																1023
Gly Glu Glu Ala Gly Ala Pro Gly Gly Gly Val Leu Ala Gly Gly Pro																
60657075																
agg gag ctg gcg gtg tgg ccg gcg gcg gca cag aga aag cgc ctc ctg																1071
Arg Glu Leu Ala Val Trp Pro Ala Ala Gln Arg Lys Arg Leu Leu																
808590																
caa ctg ccg cag tgg cgg agg cgc cgg ccg ccc gcg ccc cgc gac gac																1119
Gln Leu Pro Gln Trp Arg Arg Arg Arg Pro Pro Ala Pro Arg Asp Asp																
95100105																
ggc gag gag gcg gcc tgg gaa gaa gag tcc cct ggc ctg tca ggg ggt																1167
Gly Glu Glu Ala Ala Trp Glu Glu Glu Ser Pro Gly Leu Ser Gly Gly																
110115120																
ccg ggc ggc tcc ggg gcc gga agc acc gtg gcc gag gcc ccg ccg ggg																1215
Pro Gly Gly Ser Gly Ala Gly Ser Thr Val Ala Glu Ala Pro Pro Gly																
125130135																
acc ctg gcg ctg ctc ctg gac gaa ggc agc aag cag ctg ccg cag gcc																1263
Thr Leu Ala Leu Leu Leu Asp Glu Gly Ser Lys Gln Leu Pro Gln Ala																
140145150155																
atc atc atc gga gtg aag aag ggc ggc acg cgg gcg ctg ctg gag ttc																1311
Ile Ile Ile Gly Val Lys Lys Gly Gly Thr Arg Ala Leu Leu Glu Phe																
160165170																
ctg cgc gtg cac ccc gac gtg cgc gcc gtg ggc gcc gag ccc cac ttc																1359
Leu Arg Val His Pro Asp Val Arg Ala Val Gly Ala Glu Pro His Phe																
175180185																
ttc gac cgc agc tac gac aag ggc ctc gcc tgg tac cgg gac ctg atg																1407
Phe Asp Arg Ser Tyr Asp Lys Gly Leu Ala Trp Tyr Arg Asp Leu Met																
190195200																

ccc aga acc ctg gac ggg cag atc acc atg gag aag acg ccc agt tac 1455
 Pro Arg Thr Leu Asp Gly Gln Ile Thr Met Glu Lys Thr Pro Ser Tyr
 205 210 215

ttc gtc acg cgg gag gcc ccc gcg cgc atc tcg gcc atg tcc aag gac 1503
 Phe Val Thr Arg Glu Ala Pro Ala Arg Ile Ser Ala Met Ser Lys Asp
 220 225 230 235

acc aag ctc atc gtg gtg gtg cgg gac ccg gtg acc agg gcc atc tcg 1551
 Thr Lys Leu Ile Val Val Val Arg Asp Pro Val Thr Arg Ala Ile Ser
 240 245 250

gac tac acg cag acg ctg tcc aag cgg ccc gac atc ccc acc ttc gag 1599
 Asp Tyr Thr Gln Thr Leu Ser Lys Arg Pro Asp Ile Pro Thr Phe Glu
 255 260 265

agc ttg acg ttc aaa aac agg aca gcg ggc ctc atc gac acg tcg tgg 1647
 Ser Leu Thr Phe Lys Asn Arg Thr Ala Gly Leu Ile Asp Thr Ser Trp
 270 275 280

agc gcc atc cag atc ggc atc tac gcc aag cac ctg gag cac tgg ctg 1695
 Ser Ala Ile Gln Ile Gly Ile Tyr Ala Lys His Leu Glu His Trp Leu
 285 290 295

cgc cac ttc ccc atc cgc cag atg ctc ttc gtg agc ggc gag cgg ctc 1743
 Arg His Phe Pro Ile Arg Gln Met Leu Phe Val Ser Gly Glu Arg Leu
 300 305 310 315

atc agc gac ccg gcc ggg gag ctg ggc cgc gtg caa gac ttc ctg ggc 1791
 Ile Ser Asp Pro Ala Gly Glu Leu Gly Arg Val Gln Asp Phe Leu Gly
 320 325 330

ctc aag agg atc atc acg gac aag cac ttc tac ttc aac aag acc aag 1839
 Leu Lys Arg Ile Ile Thr Asp Lys His Phe Tyr Phe Asn Lys Thr Lys
 335 340 345

ggc ttc ccc tgc ctg aag aag gcg gag ggc agc agc cgg ccc cat tgc 1887
 Gly Phe Pro Cys Leu Lys Lys Ala Glu Gly Ser Ser Arg Pro His Cys
 350 355 360

ctg ggc aag acc aag ggc agg acc cat cct gag atc gac cgc gag gtg 1935
 Leu Gly Lys Thr Lys Gly Arg Thr His Pro Glu Ile Asp Arg Glu Val
 365 370 375

gtg cgc agg ctg cgc gag ttc tac cgg cct ttc aac ctc aag ttc tac 1983
 Val Arg Arg Leu Arg Glu Phe Tyr Arg Pro Phe Asn Leu Lys Phe Tyr
 380 385 390 395

cag atg acc ggg cac gac ttt ggc tgg gat gga taaccatata atttaaaaag 2036
 Gln Met Thr Gly His Asp Phe Gly Trp Asp Gly
 400 405

aaaaaaaaa tcaaaatata atatattttt ttaccaatcg gtagagaaga gacagtttaa 2096

tatttgtgct gaaaatatgt ttcagtatatt ttttcaatga atgttaagag attgtttctca 2156

ctcccgcccc atcttaatgt ataaccaaca ccaaacacgt ggatcaacag aaaaggaaaa 2216

tttcactcgt ctaaacactt tcaattttca gtttttattt tatgttctat ataccagtc 2276

ataaagtata agcatcagtt gtcattaaaa gttttcag 2314

<210> 8

<211> 406

<212> PRT

<213> Homo sapiens

<400> 8

Met Ala Pro Pro Gly Pro Ala Ser Ala Leu Ser Thr Ser Ala Glu Pro
1 5 10 15

Leu Ser Arg Ser Ile Phe Arg Lys Phe Leu Leu Met Leu Cys Ser Leu
20 25 30

Leu Thr Ser Leu Tyr Val Phe Tyr Cys Leu Ala Glu Arg Cys Gln Thr
35 40 45

Leu Ser Gly Pro Val Val Gly Leu Ser Gly Gly Gly Glu Glu Ala Gly
50 55 60

Ala Pro Gly Gly Gly Val Leu Ala Gly Gly Pro Arg Glu Leu Ala Val
65 70 75 80

Trp Pro Ala Ala Ala Gln Arg Lys Arg Leu Leu Gln Leu Pro Gln Trp
85 90 95

Arg Arg Arg Arg Pro Pro Ala Pro Arg Asp Asp Gly Glu Glu Ala Ala
100 105 110

Trp Glu Glu Glu Ser Pro Gly Leu Ser Gly Gly Pro Gly Gly Ser Gly
115 120 125

Ala Gly Ser Thr Val Ala Glu Ala Pro Pro Gly Thr Leu Ala Leu Leu
130 135 140

Leu Asp Glu Gly Ser Lys Gln Leu Pro Gln Ala Ile Ile Ile Gly Val
145 150 155 160

Lys Lys Gly Gly Thr Arg Ala Leu Leu Glu Phe Leu Arg Val His Pro
165 170 175

Asp Val Arg Ala Val Gly Ala Glu Pro His Phe Phe Asp Arg Ser Tyr
180 185 190

Asp Lys Gly Leu Ala Trp Tyr Arg Asp Leu Met Pro Arg Thr Leu Asp
195 200 205

Gly Gln Ile Thr Met Glu Lys Thr Pro Ser Tyr Phe Val Thr Arg Glu
210 215 220

Ala Pro Ala Arg Ile Ser Ala Met Ser Lys Asp Thr Lys Leu Ile Val
225 230 235 240

Val Val Arg Asp Pro Val Thr Arg Ala Ile Ser Asp Tyr Thr Gln Thr
 245 250 255
 Leu Ser Lys Arg Pro Asp Ile Pro Thr Phe Glu Ser Leu Thr Phe Lys
 260 265 270
 Asn Arg Thr Ala Gly Leu Ile Asp Thr Ser Trp Ser Ala Ile Gln Ile
 275 280 285
 Gly Ile Tyr Ala Lys His Leu Glu His Trp Leu Arg His Phe Pro Ile
 290 295 300
 Arg Gln Met Leu Phe Val Ser Gly Glu Arg Leu Ile Ser Asp Pro Ala
 305 310 315 320
 Gly Glu Leu Gly Arg Val Gln Asp Phe Leu Gly Leu Lys Arg Ile Ile
 325 330 335
 Thr Asp Lys His Phe Tyr Phe Asn Lys Thr Lys Gly Phe Pro Cys Leu
 340 345 350
 Lys Lys Ala Glu Gly Ser Ser Arg Pro His Cys Leu Gly Lys Thr Lys
 355 360 365
 Gly Arg Thr His Pro Glu Ile Asp Arg Glu Val Val Arg Arg Leu Arg
 370 375 380
 Glu Phe Tyr Arg Pro Phe Asn Leu Lys Phe Tyr Gln Met Thr Gly His
 385 390 395 400
 Asp Phe Gly Trp Asp Gly
 405

<210> 9
 <211> 2032
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (331)..(1500)
 <223> human 3-OST-3B

<400> 9
 gtggccaggg cgcgagagtg caacgtcctc ctggccccga gcgcgtcgtc gcgccccggg 60
 agcagaccct cgcccagcag ttaccgcgct cccgactttc cgttccagtt gcagctcctg 120
 ccgggcaaca tgtcaagagc cgccgcccgt acagctgccg ccgccacctg gggaagagca 180
 gcagcagcag cggcggccgc gggcacacgg gggcaataaa ccgagccacc cgggcgtcca 240
 gcgtgccggg gaacctcttc tgcgtcact gcccgggcggg acccacgcca tgtgctgagc 300
 catgtccctg gccgcgcccg cgggcagcgc atg ggg cag cgc ctg agt ggc ggc 354
 Met Gly Gln Arg Leu Ser Gly Gly

aga tct tgc ctc gat gtc ccc ggc cgg ctc cta ccg cag ccg ccg ccg	402
Arg Ser Cys Leu Asp Val Pro Gly Arg Leu Leu Pro Gln Pro Pro Pro	
10 15 20	
ccc ccg ccg ccg gtg agg agg aag ctc gcg ctg ctc ttc gcc atg ctc	450
Pro Pro Pro Pro Val Arg Arg Lys Leu Ala Leu Leu Phe Ala Met Leu	
25 30 35 40	
tgc gtc tgg ctc tat atg ttc ctg tac tcg tgc gcc ggc tcc tgc gcc	498
Cys Val Trp Leu Tyr Met Phe Leu Tyr Ser Cys Ala Gly Ser Cys Ala	
45 50 55	
gcc gcg ccg ggg ctg ctg ctc ctg ggc tct ggg tcc cgc gcc gca cac	546
Ala Ala Pro Gly Leu Leu Leu Leu Gly Ser Gly Ser Arg Ala Ala His	
60 65 70	
gac ccg cca gcc ctg gcc aca gct ccg gac ggg acg ccc ccc agg ctg	594
Asp Pro Pro Ala Leu Ala Thr Ala Pro Asp Gly Thr Pro Pro Arg Leu	
75 80 85	
ccg ttc ccg gcg ccg cca gcc acc cca ctg gct tca ggc aag gag atg	642
Pro Phe Arg Ala Pro Pro Ala Thr Pro Leu Ala Ser Gly Lys Glu Met	
90 95 100	
gcc gag ggc gct gcg agc ccg gag gag cag agt ccc gag gtg ccg gac	690
Ala Glu Gly Ala Ala Ser Pro Glu Glu Gln Ser Pro Glu Val Pro Asp	
105 110 115 120	
tcc cca agc ccc atc tcc agc ttt ttc agt ggg tct ggg agc aag cag	738
Ser Pro Ser Pro Ile Ser Ser Phe Phe Ser Gly Ser Gly Ser Lys Gln	
125 130 135	
ctg ccg cag gcc atc atc atc ggc gtg aag aag ggc ggc acg ccg gcg	786
Leu Pro Gln Ala Ile Ile Ile Gly Val Lys Lys Gly Gly Thr Arg Ala	
140 145 150	
ctg ctg gag ttt ctg cgc gtg cac ccc gac gtg cgc gcc gtg ggc gcc	834
Leu Leu Glu Phe Leu Arg Val His Pro Asp Val Arg Ala Val Gly Ala	
155 160 165	
gag ccc cat ttc ttc gat cgc agc tac gac aag ggc ctc gct tgg tac	882
Glu Pro His Phe Phe Asp Arg Ser Tyr Asp Lys Gly Leu Ala Trp Tyr	
170 175 180	
cgg gac ctg atg ccc aga acc ctg gac ggg cag atc acc atg gag aag	930
Arg Asp Leu Met Pro Arg Thr Leu Asp Gly Gln Ile Thr Met Glu Lys	
185 190 195 200	
acg ccc agt tac ttc gtc acg cgg gag gcc ccc gcg cgc atc tcg gcc	978
Thr Pro Ser Tyr Phe Val Thr Arg Glu Ala Pro Ala Arg Ile Ser Ala	
205 210 215	
atg tcc aag gac acc aag ctc atc gtg gtg gtg cgg gac ccg gtg acc	1026
Met Ser Lys Asp Thr Lys Leu Ile Val Val Val Arg Asp Pro Val Thr	
220 225 230	

agg gcc atc tcg gac tac acg cag acg ctg tcc aag cgg ccc gac atc 1074
 Arg Ala Ile Ser Asp Tyr Thr Gln Thr Leu Ser Lys Arg Pro Asp Ile
 235 240 245

ccc acc ttc gag agc ttg acg ttc aaa aac agg aca gcg ggc ctc atc 1122
 Pro Thr Phe Glu Ser Leu Thr Phe Lys Asn Arg Thr Ala Gly Leu Ile
 250 255 260

gac acg tcg tgg agc gcc atc cag atc ggc atc tac gcc aag cac ctg 1170
 Asp Thr Ser Trp Ser Ala Ile Gln Ile Gly Ile Tyr Ala Lys His Leu
 265 270 275 280

gag cac tgg ctg cgc cac ttc ccc atc cgc cag atg ctc ttc gtg agc 1218
 Glu His Trp Leu Arg His Phe Pro Ile Arg Gln Met Leu Phe Val Ser
 285 290 295

ggc gag cgg ctc atc agc gac ccg gcc ggg gag ctg ggc cgc gtg caa 1266
 Gly Glu Arg Leu Ile Ser Asp Pro Ala Gly Glu Leu Gly Arg Val Gln
 300 305 310

gac ttc ctg ggc ctc aag agg atc atc acg gac aag cac ttc tac ttc 1314
 Asp Phe Leu Gly Leu Lys Arg Ile Ile Thr Asp Lys His Phe Tyr Phe
 315 320 325

aac aag acc aag ggc ttc ccc tgc ctg aag aag gcg gag ggc agc agc 1362
 Asn Lys Thr Lys Gly Phe Pro Cys Leu Lys Lys Ala Glu Gly Ser Ser
 330 335 340

cgg ccc cat tgc ctg ggc aag acc aag ggc agg acc cat cct gag atc 1410
 Arg Pro His Cys Leu Gly Lys Thr Lys Gly Arg Thr His Pro Glu Ile
 345 350 355 360

gac cgc gag gtg gtg cgc agg ctg cgc gag ttc tac cgg cct ttc aac 1458
 Asp Arg Glu Val Val Arg Arg Leu Arg Glu Phe Tyr Arg Pro Phe Asn
 365 370 375

ctc aag ttc tac cag atg acc ggg cac gac ttt ggc tgg gat 1500
 Leu Lys Phe Tyr Gln Met Thr Gly His Asp Phe Gly Trp Asp
 380 385 390

tgagcagacc cgggctatgt accttaccga cgtggcttat ctattgacag agattatatg 1560

tatgtaaaat gtacagaaat ctattttata ataatttatt tttaattcat aagcaattaa 1620

ttcactaagc tgcctagcca cactcttttag agagtttagct tcataatctg ttaacattcc 1680

aaagtgttta actctagtat ttcgttttct tcttcacaat tgatggtgct tctatttttt 1740

cttctcccct acctgttata tttaaaacaa agaaaagcac aacttgagat ttttggtggt 1800

acgggtattc agccttcagt caccgtctga gttctccagt tgctgcctcc ttgtcttgct 1860

ttgggtctcc cattccagct tccctgtctc ttcctgctg tgtacctcgt aggaacgctg 1920

agctgcctca acagggtgtg attctgaagg gcaggcctca tgcagcagcc tccttgacga 1980

tgtgggtgtcc cgtccaatga tgtagcctga aagccacagc cctaggggttc tg

2032

<210> 10

<211> 390

<212> PRT

<213> Homo sapiens

<400> 10

Met Gly Gln Arg Leu Ser Gly Gly Arg Ser Cys Leu Asp Val Pro Gly
1 5 10 15

Arg Leu Leu Pro Gln Pro Pro Pro Pro Pro Pro Val Arg Arg Lys
20 25 30

Leu Ala Leu Leu Phe Ala Met Leu Cys Val Trp Leu Tyr Met Phe Leu
35 40 45

Tyr Ser Cys Ala Gly Ser Cys Ala Ala Ala Pro Gly Leu Leu Leu Leu
50 55 60

Gly Ser Gly Ser Arg Ala Ala His Asp Pro Pro Ala Leu Ala Thr Ala
65 70 75 80

Pro Asp Gly Thr Pro Pro Arg Leu Pro Phe Arg Ala Pro Pro Ala Thr
85 90 95

Pro Leu Ala Ser Gly Lys Glu Met Ala Glu Gly Ala Ala Ser Pro Glu
100 105 110

Glu Gln Ser Pro Glu Val Pro Asp Ser Pro Ser Pro Ile Ser Ser Phe
115 120 125

Phe Ser Gly Ser Gly Ser Lys Gln Leu Pro Gln Ala Ile Ile Ile Gly
130 135 140

Val Lys Lys Gly Gly Thr Arg Ala Leu Leu Glu Phe Leu Arg Val His
145 150 155 160

Pro Asp Val Arg Ala Val Gly Ala Glu Pro His Phe Phe Asp Arg Ser
165 170 175

Tyr Asp Lys Gly Leu Ala Trp Tyr Arg Asp Leu Met Pro Arg Thr Leu
180 185 190

Asp Gly Gln Ile Thr Met Glu Lys Thr Pro Ser Tyr Phe Val Thr Arg
195 200 205

Glu Ala Pro Ala Arg Ile Ser Ala Met Ser Lys Asp Thr Lys Leu Ile
210 215 220

Val Val Val Arg Asp Pro Val Thr Arg Ala Ile Ser Asp Tyr Thr Gln
225 230 235 240

Thr Leu Ser Lys Arg Pro Asp Ile Pro Thr Phe Glu Ser Leu Thr Phe
245 250 255

Lys Asn Arg Thr Ala Gly Leu Ile Asp Thr Ser Trp Ser Ala Ile Gln
 260 265 270
 Ile Gly Ile Tyr Ala Lys His Leu Glu His Trp Leu Arg His Phe Pro
 275 280 285
 Ile Arg Gln Met Leu Phe Val Ser Gly Glu Arg Leu Ile Ser Asp Pro
 290 295 300
 Ala Gly Glu Leu Gly Arg Val Gln Asp Phe Leu Gly Leu Lys Arg Ile
 305 310 315 320
 Ile Thr Asp Lys His Phe Tyr Phe Asn Lys Thr Lys Gly Phe Pro Cys
 325 330 335
 Leu Lys Lys Ala Glu Gly Ser Ser Arg Pro His Cys Leu Gly Lys Thr
 340 345 350
 Lys Gly Arg Thr His Pro Glu Ile Asp Arg Glu Val Val Arg Arg Leu
 355 360 365
 Arg Glu Phe Tyr Arg Pro Phe Asn Leu Lys Phe Tyr Gln Met Thr Gly
 370 375 380
 His Asp Phe Gly Trp Asp
 385 390

<210> 11
 <211> 3658
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (847) .. (2214)

<220>
 <223> Predicted human 3-OST-4 hnRNA

<400> 11
 gaggatatcc cgggcgagag aaggagggt cggggatggg ctgagttgga gtcccagagg 60
 aaaagcggaa gcgagagctt cgtcaccgc tgtcttcag ctcccggtgc gcggcaccgg 120
 aggcaggcgt tgggctttac ctctctaaaa gtactggggc aaaggaatgg agaacacggc 180
 gtcccagagct cccaagggag gggagtaaac gaggtggggg ggggaacacc ccaagtgcgt 240
 gcgtgctggg gggctggggg gcaagatctc cgttctcccg ggtgccccag cctagcgca 300
 cgctccgct ccccgcccc ctctgcaggc gcgcgcgagg cgcaccccc ttccctcggc 360
 ggcgcggggc gcgcgcccgg cccctctctc ctccctccg cgctctctct ctctcccggc 420
 agaaaagttag cagcggggaa ggaactctgg gctgcaacag cgcgcggcgg cggcggcaga 480

ggctgaagca gaagccgcgg cggagccggg gaagcggggg cgctgcagac ggagcaggtg 540
 ccgccggcgg gtccgcgcgc cccctcggt ccccttgctt gaggtgagg ggggggcggg 600
 ggtggggggg ccactcggac tcggcgggca gcgtggggcg gggggccatg cggccgggct 660
 ccccccctggc gcagcgggac agcggccagg gccggggggc cagcggcgtc gcttcatgca 720
 gccggggcgg ctgggcagcg gcggcggcgg cggcggcggc ggccggcggg gcggcggctg 780
 aaaccatgtc cgggcagcgc cgggggctgc cgccgcgcgc gccgcgcgc cgagccggga 840
 gccgcg atg gcc cgg tgg ccc gca cct cct ccg cct ccg cct ccg cct 888
 Met Ala Arg Trp Pro Ala Pro Pro Pro Pro Pro Pro Pro
 1 5 10
 cca cct ctg gcc gcg ccg ccg ccg ccc ggc gcc tct gct aag ggg ccg 936
 Pro Pro Leu Ala Ala Pro Pro Pro Pro Gly Ala Ser Ala Lys Gly Pro
 15 20 25 30
 ccg gcg cgc aag ctg ctt ttt atg tgc acc ttg tcc ctg tct gtc acc 984
 Pro Ala Arg Lys Leu Leu Phe Met Cys Thr Leu Ser Leu Ser Val Thr
 35 40 45
 tac ctg tgc tac agc ctc ctg ggc ggc tgc ggc tcc ctg caa ttc cct 1032
 Tyr Leu Cys Tyr Ser Leu Leu Gly Gly Ser Gly Ser Leu Gln Phe Pro
 50 55 60
 ctg gcg ctg cag gag tgc ccg ggc gcc gcc gcc gag ccc ccg ccg agc 1080
 Leu Ala Leu Gln Glu Ser Pro Gly Ala Ala Ala Glu Pro Pro Pro Ser
 65 70 75
 ccg ccg cca ccc tct ctg ctg cct acc ccc gtg cgc ctc ggc gcc ccc 1128
 Pro Pro Pro Pro Ser Leu Leu Pro Thr Pro Val Arg Leu Gly Ala Pro
 80 85 90
 tgc cag ccg ccc gcg ccg ccg ccg ctg gac aac gcg agc cac ggg gag 1176
 Ser Gln Pro Pro Ala Pro Pro Pro Leu Asp Asn Ala Ser His Gly Glu
 95 100 105 110
 ccg ccc gag ccc cca gag cag cca gcc gcc ccc ggg acc gac ggc tgg 1224
 Pro Pro Glu Pro Pro Glu Gln Pro Ala Ala Pro Gly Thr Asp Gly Trp
 115 120 125
 ggg ctg ccg agc ggc ggc gga ggc gcc cgg gac gcc tgg ctc cgg acc 1272
 Gly Leu Pro Ser Gly Gly Gly Gly Ala Arg Asp Ala Trp Leu Arg Thr
 130 135 140
 ccg ctg gcc ccc agc gag atg atc acg gct cag agc gcg ctg ccg gag 1320
 Pro Leu Ala Pro Ser Glu Met Ile Thr Ala Gln Ser Ala Leu Pro Glu
 145 150 155
 agg gaa gcg cag gag tcc agc acc acc gac gag gat ctc gca ggc cgg 1368
 Arg Glu Ala Gln Glu Ser Ser Thr Thr Asp Glu Asp Leu Ala Gly Arg
 160 165 170
 aga gcg gcc aac ggg agc agc gag agg ggc ggc gcc gtc agc acc ccc 1416

Arg	Ala	Ala	Asn	Gly	Ser	Ser	Glu	Arg	Gly	Gly	Ala	Val	Ser	Thr	Pro	
175					180					185					190	
gac	tat	ggg	gag	aag	aag	ctg	cca	cag	gcg	ctc	atc	atc	ggg	gtc	aag	1464
Asp	Tyr	Gly	Glu	Lys	Lys	Leu	Pro	Gln	Ala	Leu	Ile	Ile	Gly	Val	Lys	
				195					200					205		
aaa	gga	ggg	acc	cgc	gcg	ctg	ctg	gag	gcg	atc	cgc	gtg	cac	ccg	gac	1512
Lys	Gly	Gly	Thr	Arg	Ala	Leu	Leu	Glu	Ala	Ile	Arg	Val	His	Pro	Asp	
			210					215				220				
gtg	cgg	gcg	gtg	ggc	gta	gag	ccg	cac	ttc	ttc	gac	agg	aac	tac	gaa	1560
Val	Arg	Ala	Val	Gly	Val	Glu	Pro	His	Phe	Phe	Asp	Arg	Asn	Tyr	Glu	
		225					230					235				
aag	ggg	ttg	gag	tgg	tac	aga	aat	gtg	atg	ccc	aag	act	ttg	gat	ggg	1608
Lys	Gly	Leu	Glu	Trp	Tyr	Arg	Asn	Val	Met	Pro	Lys	Thr	Leu	Asp	Gly	
	240					245					250					
caa	ata	acc	atg	gag	aag	act	cca	agt	tac	ttt	gtg	aca	aat	gag	gct	1656
Gln	Ile	Thr	Met	Glu	Lys	Thr	Pro	Ser	Tyr	Phe	Val	Thr	Asn	Glu	Ala	
255					260					265					270	
ccc	aag	cgc	att	cac	tcc	atg	gcc	aag	gac	atc	aaa	ctg	att	gtg	gtg	1704
Pro	Lys	Arg	Ile	His	Ser	Met	Ala	Lys	Asp	Ile	Lys	Leu	Ile	Val	Val	
				275					280					285		
gtg	aga	aac	ccc	gtg	acc	agg	gcc	atc	tct	gac	tac	acg	cag	aca	ctg	1752
Val	Arg	Asn	Pro	Val	Thr	Arg	Ala	Ile	Ser	Asp	Tyr	Thr	Gln	Thr	Leu	
			290					295					300			
tca	aag	aaa	ccc	gag	atc	ccc	acc	ttt	gag	gtg	ctg	gcc	ttc	aaa	aac	1800
Ser	Lys	Lys	Pro	Glu	Ile	Pro	Thr	Phe	Glu	Val	Leu	Ala	Phe	Lys	Asn	
		305					310					315				
cgg	acc	ctc	ggg	ctg	atc	gat	gct	tcc	tgg	agt	gcc	att	cga	ata	ggg	1848
Arg	Thr	Leu	Gly	Leu	Ile	Asp	Ala	Ser	Trp	Ser	Ala	Ile	Arg	Ile	Gly	
	320					325					330					
atc	tat	gcg	ctg	cat	ctg	gaa	aac	tgg	ctc	cag	tat	ttc	ccc	ctc	tcc	1896
Ile	Tyr	Ala	Leu	His	Leu	Glu	Asn	Trp	Leu	Gln	Tyr	Phe	Pro	Leu	Ser	
335					340					345					350	
cag	atc	ctc	ttt	gtc	agt	ggt	gag	cga	ctc	att	gtg	gac	ccc	gcc	ggg	1944
Gln	Ile	Leu	Phe	Val	Ser	Gly	Glu	Arg	Leu	Ile	Val	Asp	Pro	Ala	Gly	
				355					360					365		
gaa	atg	gcc	aaa	gta	cag	gat	ttt	cta	ggc	ctc	aaa	cgt	gtt	gtg	act	1992
Glu	Met	Ala	Lys	Val	Gln	Asp	Phe	Leu	Gly	Leu	Lys	Arg	Val	Val	Thr	
			370					375					380			
aag	aag	cat	ttc	tat	ttc	aac	aaa	acc	aag	ggg	ttc	cct	tgc	cta	aag	2040
Lys	Lys	His	Phe	Tyr	Phe	Asn	Lys	Thr	Lys	Gly	Phe	Pro	Cys	Leu	Lys	
		385					390					395				
aag	cca	gaa	gac	agc	agt	gcc	ccg	agg	tgc	tta	ggc	aag	agc	aaa	ggt	2088
Lys	Pro	Glu	Asp	Ser	Ser	Ala	Pro	Arg	Cys	Leu	Gly	Lys	Ser	Lys	Gly	

400

405

410

cgg act cat cct cgc att gac cca gat gtc atc cac aga ctg agg aaa 2136
 Arg Thr His Pro Arg Ile Asp Pro Asp Val Ile His Arg Leu Arg Lys
 415 420 425 430

ttc tac aaa ccc ttc aac ttg atg ttt tac caa atg act ggt caa gat 2184
 Phe Tyr Lys Pro Phe Asn Leu Met Phe Tyr Gln Met Thr Gly Gln Asp
 435 440 445

ttt cag tgg gaa cag gaa gag ggt gat aaa tgaggctaga gaggcagagg 2234
 Phe Gln Trp Glu Gln Glu Glu Gly Asp Lys
 450 455

aaggctagtc aataagctaa ggaggctcct tgccctgagtc cttgaatacc ccagcttctg 2294

cagcttcact tgctggagtg ccaagtagat ctccctcctcc ttcctgcagc caggattgcc 2354

tccagtgcctg ttagcttagg caaacagggtg gatcccatgg catcccatg gaggaaccag 2414

gcccatctgg gcagcagcat ctgggtgacc agatggccac cagaaccac tgttcattct 2474

tatcttctgc tagttaatat agcctgaaga cagaggataa atagttgtca atgtcagaga 2534

cagtgcctatt aatgtatatg tgagcgacaa aaaaggctctg ctttataggg gttctcactc 2594

tagcttgggg agcccagggt tctagccctg tatctgtcat gggcacctgc tgtctaaacc 2654

tctgcttggg cttctcccca gaatgcactt tgtggctgag tgctccagga ctccctaggga 2714

gcaagctcct cctctaaagg tgtttctagt cttctcttta aaggctctcat ccacacaacc 2774

ctgacttctt cctcccccac atcatgaagg cagaggcatg cacattcctc actgaaaaag 2834

aaaacacaca cccaccacaca cacacacaca cagaagaaaa tgaaagctga cacacctcga 2894

agccttcttt ccaagagccc tctaaatggg gttgggtctc actcttcatg agtatcctgg 2954

gttgtgcaga agcttagcat atgcccttgt gttcggatca ggcccacagg gctgctcaaa 3014

gagtagagta attgtaaccg aggtcagagc tctgggggtg gcagagatga gtggccatat 3074

ctgggggtaa aagaagaaat cctgtcctct tgggtgggagg ttaccttacc tgaagaccat 3134

ctctcccaag cactgtagtt ctgagcatgt ttttgggggtg gactctgtcc cctaggggtcc 3194

ctagaagggc aaagaccaga gagttgacaa gtctgttatt aggaataatc cttagccatg 3254

taatggagaa aggagcagtc agcattcttc caatttgccc caccaccacc tcctcgggct 3314

tcattttctc tatntagaga tggcagagag tgaggtagtg gcgagaaagc tgactccatt 3374

catcagatcc agtttatgag ggttgggggt gagcaagggc tgtctgcaga aacccccatc 3434

aagagctgct gaatgaagtg tcccttccca tcagtttgat tcaattaaaa tgcattcattt 3494

gacataaagc acttggtcac agatctccaa aaccaggaat tgttctagta aaactggaaa 3554

tttgtatgag tgggggggagt taaatctggt cagctgttat taaactgtca tttctccgcg 3614
 taaatgaaaa cctgtgttgtt ataaagctta atgcaacctg atta 3658

<210> 12
 <211> 456
 <212> PRT
 <213> Homo sapiens

<400> 12
 Met Ala Arg Trp Pro Ala Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro
 1 5 10 15
 Leu Ala Ala Pro Pro Pro Pro Gly Ala Ser Ala Lys Gly Pro Pro Ala
 20 25 30
 Arg Lys Leu Leu Phe Met Cys Thr Leu Ser Leu Ser Val Thr Tyr Leu
 35 40 45
 Cys Tyr Ser Leu Leu Gly Gly Ser Gly Ser Leu Gln Phe Pro Leu Ala
 50 55 60
 Leu Gln Glu Ser Pro Gly Ala Ala Ala Glu Pro Pro Pro Ser Pro Pro
 65 70 75 80
 Pro Pro Ser Leu Leu Pro Thr Pro Val Arg Leu Gly Ala Pro Ser Gln
 85 90 95
 Pro Pro Ala Pro Pro Pro Leu Asp Asn Ala Ser His Gly Glu Pro Pro
 100 105 110
 Glu Pro Pro Glu Gln Pro Ala Ala Pro Gly Thr Asp Gly Trp Gly Leu
 115 120 125
 Pro Ser Gly Gly Gly Gly Ala Arg Asp Ala Trp Leu Arg Thr Pro Leu
 130 135 140
 Ala Pro Ser Glu Met Ile Thr Ala Gln Ser Ala Leu Pro Glu Arg Glu
 145 150 155 160
 Ala Gln Glu Ser Ser Thr Thr Asp Glu Asp Leu Ala Gly Arg Arg Ala
 165 170 175
 Ala Asn Gly Ser Ser Glu Arg Gly Gly Ala Val Ser Thr Pro Asp Tyr
 180 185 190
 Gly Glu Lys Lys Leu Pro Gln Ala Leu Ile Ile Gly Val Lys Lys Gly
 195 200 205
 Gly Thr Arg Ala Leu Leu Glu Ala Ile Arg Val His Pro Asp Val Arg
 210 215 220
 Ala Val Gly Val Glu Pro His Phe Phe Asp Arg Asn Tyr Glu Lys Gly
 225 230 235 240

Leu Glu Trp Tyr Arg Asn Val Met Pro Lys Thr Leu Asp Gly Gln Ile
 245 250 255
 Thr Met Glu Lys Thr Pro Ser Tyr Phe Val Thr Asn Glu Ala Pro Lys
 260 265 270
 Arg Ile His Ser Met Ala Lys Asp Ile Lys Leu Ile Val Val Val Arg
 275 280 285
 Asn Pro Val Thr Arg Ala Ile Ser Asp Tyr Thr Gln Thr Leu Ser Lys
 290 295 300
 Lys Pro Glu Ile Pro Thr Phe Glu Val Leu Ala Phe Lys Asn Arg Thr
 305 310 315 320
 Leu Gly Leu Ile Asp Ala Ser Trp Ser Ala Ile Arg Ile Gly Ile Tyr
 325 330 335
 Ala Leu His Leu Glu Asn Trp Leu Gln Tyr Phe Pro Leu Ser Gln Ile
 340 345 350
 Leu Phe Val Ser Gly Glu Arg Leu Ile Val Asp Pro Ala Gly Glu Met
 355 360 365
 Ala Lys Val Gln Asp Phe Leu Gly Leu Lys Arg Val Val Thr Lys Lys
 370 375 380
 His Phe Tyr Phe Asn Lys Thr Lys Gly Phe Pro Cys Leu Lys Lys Pro
 385 390 395 400
 Glu Asp Ser Ser Ala Pro Arg Cys Leu Gly Lys Ser Lys Gly Arg Thr
 405 410 415
 His Pro Arg Ile Asp Pro Asp Val Ile His Arg Leu Arg Lys Phe Tyr
 420 425 430
 Lys Pro Phe Asn Leu Met Phe Tyr Gln Met Thr Gly Gln Asp Phe Gln
 435 440 445
 Trp Glu Gln Glu Glu Gly Asp Lys
 450 455

<210> 13
 <211> 284
 <212> PRT
 <213> Homo sapiens

<220>
 <223> human NST-1 (aa 599 to 882)

<400> 13
 Lys Thr Cys Asp Arg Phe Pro Lys Leu Leu Ile Ile Gly Pro Gln Lys
 1 5 10 15
 Thr Gly Thr Thr Ala Leu Tyr Leu Phe Leu Gly Met His Pro Asp Leu
 20 25 30

<400> 14

Lys	Thr	Cys	Asp	Arg	Leu	Pro	Lys	Phe	Leu	Ile	Val	Gly	Pro	Gln	Lys
1				5					10					15	
Thr	Gly	Thr	Thr	Ala	Ile	His	Phe	Phe	Leu	Ser	Leu	His	Pro	Ala	Val
			20					25					30		
Thr	Ser	Ser	Phe	Pro	Ser	Pro	Ser	Thr	Phe	Glu	Glu	Ile	Gln	Phe	Phe
		35					40					45			
Asn	Ser	Pro	Asn	Tyr	His	Lys	Gly	Ile	Asp	Trp	Tyr	Met	Asp	Phe	Phe
	50					55					60				
Pro	Val	Pro	Ser	Asn	Ala	Ser	Thr	Asp	Phe	Leu	Phe	Glu	Lys	Ser	Ala
65					70					75					80
Thr	Tyr	Phe	Asp	Ser	Glu	Val	Val	Pro	Arg	Arg	Gly	Ala	Ala	Leu	Leu
				85					90					95	
Pro	Arg	Ala	Lys	Ile	Ile	Thr	Val	Leu	Thr	Asn	Pro	Ala	Asp	Arg	Ala
			100					105					110		
Tyr	Ser	Trp	Tyr	Gln	His	Gln	Arg	Ala	His	Gly	Asp	Pro	Val	Ala	Leu
	115						120					125			
Asn	Tyr	Thr	Phe	Tyr	Gln	Val	Ile	Ser	Ala	Ser	Ser	Gln	Thr	Pro	Leu
	130					135						140			
Ala	Leu	Arg	Ser	Leu	Gln	Asn	Arg	Cys	Leu	Val	Pro	Gly	Tyr	Tyr	Ser
145					150					155					160
Thr	His	Leu	Gln	Arg	Trp	Leu	Thr	Tyr	Tyr	Pro	Ser	Gly	Gln	Leu	Leu
				165					170					175	
Ile	Val	Asp	Gly	Gln	Glu	Leu	Arg	Thr	Asn	Pro	Ala	Ala	Ser	Met	Glu
			180					185					190		
Ser	Ile	Gln	Lys	Phe	Leu	Gly	Ile	Thr	Pro	Phe	Leu	Asn	Tyr	Thr	Arg
		195					200					205			
Thr	Leu	Arg	Phe	Asp	Asp	Asp	Lys	Gly	Phe	Trp	Cys	Gln	Gly	Leu	Glu
	210					215					220				
Gly	Gly	Lys	Thr	Arg	Cys	Leu	Gly	Arg	Ser	Lys	Gly	Arg	Arg	Tyr	Pro
225					230					235					240
Asp	Met	Asp	Thr	Glu	Ser	Arg	Leu	Phe	Leu	Thr	Asp	Phe	Phe	Arg	Asn
				245					250					255	
His	Asn	Leu	Glu	Leu	Ser	Lys	Leu	Leu	Ser	Arg	Leu	Gly	Gln	Pro	Val
		260						265					270		
Pro	Ser	Trp	Leu	Arg	Glu	Glu	Leu	Gln	His	Ser	Ser	Leu	Gly		
		275					280						285		

<210> 15
<211> 291
<212> PRT
<213> Caenorhabditis elegans

<220>
<223> putative C. elegans 3-OST

<400> 15

Met Lys Tyr Arg Leu Leu Ile Leu His Leu Ile Asp Leu Ile Ser
1 5 10 15

Cys Gly Val Ile Pro Asn Thr Ser Lys Lys Arg Phe Pro Asp Ala Ile
20 25 30

Ile Val Gly Val Lys Lys Ser Gly Thr Arg Ala Leu Leu Glu Phe Leu
35 40 45

Arg Val Asn Pro Leu Ile Lys Ala Pro Gly Pro Glu Val His Phe Phe
50 55 60

Asp Lys Asn Phe Asn Lys Gly Leu Glu Trp Tyr Arg Glu Gln Met Pro
65 70 75 80

Glu Thr Lys Phe Gly Glu Val Thr Ile Glu Lys Ser Pro Ala Tyr Phe
85 90 95

His Ser Lys Met Ala Pro Glu Arg Ile Lys Ser Leu Asn Pro Asn Thr
100 105 110

Lys Ile Ile Ile Val Val Arg Asp Pro Val Thr Arg Ala Ile Ser Asp
115 120 125

Tyr Thr Gln Ser Ser Ser Lys Arg Lys Arg Val Gly Leu Met Pro Ser
130 135 140

Phe Glu Thr Met Ala Val Gly Asn Cys Ala Asn Trp Leu Arg Thr Asn
145 150 155 160

Cys Thr Thr Lys Thr Arg Gly Val Asn Ala Gly Trp Gly Ala Ile Arg
165 170 175

Ile Gly Val Tyr His Lys His Met Lys Arg Trp Leu Asp His Phe Pro
180 185 190

Ile Glu Asn Ile His Ile Val Asp Gly Glu Lys Leu Ile Ser Asn Pro
195 200 205

Ala Asp Glu Ile Ser Ala Thr Glu Lys Phe Leu Gly Leu Lys Pro Val
210 215 220

Ala Lys Pro Glu Lys Phe Gly Val Asp Pro Ile Lys Lys Phe Pro Cys
225 230 235 240

Ile Lys Asn Glu Asp Gly Lys Leu His Cys Leu Gly Lys Thr Lys Gly
245 250 255

Arg His His Pro Asp Val Glu Pro Ser Val Leu Lys Thr Leu Arg Glu
260 265 270

Phe Tyr Gly Pro Glu Asn Lys Lys Phe Tyr Gln Met Ile Asn His Trp
275 280 285

Phe Asp Trp
290

<210> 16
<211> 4045
<212> DNA
<213> Homo sapiens

<220>
<223> 3-OST-4 5' promoter/exon

<400> 16
gaattctgtg ggtgttggca ggggagacag aaaactatct tccatcgagt cttcggatcc 60
attgggaatg cctggatgac gtcagagttc gccctgtgta ggtagctccc acttttcatt 120
gtaggtttct caaggacttg ctctagaaa aagcgtggct caaaagtaga taaaaaatag 180
gcaactgcct aagtgtgaaa ttacaaaagt tcctctccaa aaaagcccg ctcctcccta 240
tcacttgtgg gcctgacatt ttaccaaagg ggctctatc tttcaagagt ttgttattaa 300
agcgtgacta tttgaggatt ggaggcaaaa gggatactga gaaatgtcct tactagcagt 360
gtcaaggcaa gtgacataaa tgtgtggggg ggcaacttgt atgagcactg tgaaaacggc 420
agcatgttca ctctacttct cagctctgac tgaggggctc aaagttcagg atctgctgat 480
ttttcaacag taacgtcctc tccaaggtgt tttttttttt tccttttttg ggaaagcccc 540
cagtttaaac tattgcagcc agtttacatt tcttaatgtc actgtgctgg ccacattcag 600
agctccatth gccaccatcg gttttgatac ctttttacca aaacctttcg aaatttgaga 660
gcccatcttt agtaaaactg ggcattggagc agattcgtht ggattgctga gaggggagat 720
agaaaagtht ggggtgctagg caggaaactgc aaggaggacc tgggccatat gccagacatc 780
tagtgcttgg gccttgaaag ggagactggg cgctgacaag gcaatatctg ttgcaacca 840
ggcttcctag atgaccacct tggatcatgg ctcgagcac agggagggct gggcagtgtc 900
tgtgtttctc tccgttccag ttggccccct cccattgaca ttacagtaat gcagttgtgt 960
gctgtttgaa aaagcatccc tagttacaca gaatgattta caggacacca gactctgcat 1020
ttcagagggtc tccagtgtac cataaaaaat atattataaa agaataatct ttatctgaac 1080
taaagctgca gtgaaggaaa ctctgtgtcca gctgagagca gcagtgtgct tttgttcact 1140

cagggaaaaag tccgtgttct ttatcttatt tgattaacat tttttttctc tttggcacta 1200
ggtatctcgt taatatctag acattatata ctttttcttt caactgggtt tcttattacg 1260
tgatcaaaca aaaccagaga tgccagccac agcagccaac aagggaaaaag cagccccgtc 1320
aggcaccacg ctgtggctgg ggggcaaagt gtactcacta gagccacccc caggaggcac 1380
ctggcagagc tctgtgcaga gccagccccg gttgcagaaa gctgagtttg ttggagtgcc 1440
tcagttgatc actctgtctc tttctcccat ttcctcact tccctgagca aaatgcaaca 1500
ggaagcaaag tctagttgtg aatcttccaa agccttctga tgtttaccat gttccccag 1560
gagagggagg tgaggggtgg agatctctct gcaaagaaaa tacacttaa aaatttcagc 1620
gagccgatgc acagacaccc agcaaccacg cttgtctccg cttattaggt gttcagagcg 1680
acagtgggcc cacactatct cagtccagga aacctgaac tccgttagtg gcaatgcccc 1740
cgaagaggcg caggtgtgtg cacctgtgat taaggggtgc gaggaggggc agcctcatct 1800
cttgaagcag aaagtgttgt cacctggtga tgggacagag ggaaaagctc tggggctggg 1860
aaacctgggg gcttgtgtca aagctccacc catcaggagc ttcaagagaa gatggggggc 1920
ggggggcggt ggctggaaag atggaagttg ggatgggaaa gcggtttagt aaaaggattc 1980
actcctggac cgaaggcagg aggatatccc gggcgagaga agggaggggc ggggatgggc 2040
tgagttggag tcccagagga aaagcggaag cgagagcttc gtcacccgct gtcttcacg 2100
tcccgggtgc cggcaccgga ggcaggcggt gggctttacc tctctaaaag tactggggca 2160
aaggaatgga gaacacggcg tcccagctc ccaagggagg ggagtaaag aggtgggggtg 2220
gggaacaccc caagtgcgtg cgtgctgggg ggctgggggg cacgatctcc gttctcccg 2280
gtgccccagc cctagcgac gcctccgctc ccccgcccc ttgcaggcg cgcgcaggc 2340
gcacccccct tccctcgcg gcgcggggcg cgcgcgggc cccctctcc tccctccgc 2400
gcctctctc tctcccgca gaaagttagc agcggggaag gaactctggg ctgcaacagc 2460
gcgcggcggc ggcggcagag gctgaagcag aagccgcggc ggagccgggg aagcgggggc 2520
gctgcagacg gagcaggtgc cgcggcggg tccgcgcgc cccctcggtc cccttgctg 2580
aggctgaggg gggggcggtg gtgggggggc cactcggact cggcgggcag cgtggggcg 2640
ggggccatgc ggccgggctc cccctggcg cagcgggaca gcggccaggg ccgggggcgc 2700
agcggcgctc ctcatgcag ccggggcggc tgggcagcgg cggcggcggc ggcgcggcg 2760
gcggcggggg cgcgggctga aacctgtcc gggcagcgcc gggggctgcc gccgcgcgc 2820
ccgcgcgcgc gagcggggag ccgcgatggc ccggtggccc gcacctctc cgcctccgc 2880

tccgcctcca cctctggcgg cgccgcgcgc gcccgggcgc tctgctaagg ggccgcccgc 2940
 gcgcaagctg ctttttatgt gcaccttgtc cctgtctgtc acctacctgt gctacagcct 3000
 cctggggcggc tcgggctccc tgcaattccc tctggcgctg caggagtcgc cgggcgcgcg 3060
 cgccgagccc ccgcgagacc cgccgcacac ctctctgctg cctacccccg tgcgcctcgg 3120
 cgccccctcg cagccgcccc cgccgcgcgc gctggacaac gcgagccacg gggagccgcc 3180
 cgagccccca gagcagccag ccgcccccgg gaccgacggc tgggggctgc cgagcggcgg 3240
 cggagggcgc cgggacgcct ggctccggac cccgctggcc cccagcgaga tgatcacggc 3300
 tcagagcgcg ctgccggaga gggaagcgca ggagtccagc accaccgacg aggatctcgc 3360
 aggcgggaga gcggccaacg ggagcagcga gaggggcggc gccgtcagca cccccgacta 3420
 tggggagaag aagctgccac aggcgctcat catcggggtc aagaaaggag ggaccgcgcg 3480
 gctgctggag gcgatccgcg tgcacccgga cgtgcggggc gtgggcgtag agccgcactt 3540
 cttcgacagg aactacgaaa aggggttgga gtggtacagg taggacctg ggctccgcgg 3600
 gctggtggag acgcgtgggg gagacgcgga ggggaagccg cggctttcca cgcccttcga 3660
 gcatccaggc accgtcccga gaggcccaag ccccgcgag ggctctgcaa accctggcgg 3720
 cgttgctcag ggggatcggc tgagagggct ggactccagc gaaaggtcac tttatttcag 3780
 ggcgagggga ggaggtgtca ccctgcctg cctccgcgc tcctcatcca aggaggtgct 3840
 gtctgaatct gccagctcc aagcctggga acccccagcc ctctgcctg ctgggtgttt 3900
 ccgaaaccag gctcttgccg ggttctggga ttctgggcag aggactttga ggagtgcagc 3960
 aggatggcta aattgactaa ggggatttga ggtcccctgg aatctcttaa aatcacccctc 4020
 aaacgcattt gcgtggctgg aattc 4045